Shape Change Along Geodesics with Application to Cleft Lip Surgery

Julian J. Faraway*and Carroll-Ann Trotman†

April 28, 2011

Abstract

Continuous shape change is represented as curves in the shape space. A method for checking the closeness of these curves to a geodesic is presented. Three large databases of short human motions are considered and shown to be well-approximated by geodesics. The motions are thus approximated by two shapes on the geodesic and the rate of progress along the path. An analysis of facial motion data taken from a study of subjects with cleft lip/palate is presented that allows the motion to be considered independently from the static shape. Inferential methods for assessing the change in motion are presented. The construction of predicted animated motions is discussed.

Keywords: facial motion, functional data analysis, geodesics, landmarks, principal component analysis, shape analysis

1 Introduction

Shapes can be described by a set of landmarks and can change in two different ways. The shape as a whole can translate and rotate and possibly change in overall size or the internal relationship between the landmarks can change. Our application concerns facial soft tissue motion — thus we are interested in the motion of the face, but not the motion of the head. Hence, we are interested only in the change in the shape of the landmarks on the face relative to one another and not in rotations or translations of the landmarks as a whole that would be caused by head motion. Performing statistics on shape data is not easy because shapes lie in a nonlinear space making even simple operations like averaging difficult. Introductions to shape statistics can be found in Small (1996), Dryden and Mardia (1998), Bookstein (1991) and Kendall et al. (1999)

Facial motion produces shapes that change continuously in time. Consider a three dimensional shape specified by k landmarks changing in time as $Y(t)$, which can be viewed as a particular kind

*Department of Mathematical Sciences, University of Bath, BA2 7AY, United Kingdom, jjf23@bath.ac.uk
†University of Maryland and University of North Carolina

of $3k$-dimensional function of t. In practice, $Y(t)$ will be recorded on a relatively dense grid of timepoints . We are interested in collections of n shape curves, $Y_1(t), \ldots, Y_n(t)$ and how they relate to covariate information in vectors x_i for $i = 1, \ldots, n$. For example, we might want to investigate the differences between two groups or predict how the shape curve will change as a given predictor varies. This can be viewed as shape-valued functional data analysis.

Shape changes are represented by curves in the nonlinear shape space (formally Σ_3^k as described in the references above). A geodesic is the shortest curve between two points in this space. We shall show that the shape curves in our particular data are well approximated by geodesics. We also show that the shape curves in two other large databases of short human motions are well approximated by geodesics. Thus a single motion is approximated by a geodesic (which can be described by its endpoints) and the rate of progress along the geodesic. With this reduction in dimensionality, modelling becomes substantially easier. Thus we make no claim of a universal approach to modelling this type of data, rather we present a method that will be effective for some datasets. In particular, it is only likely to be effective where the variation in the shape space is relatively small.

We present the context behind our application and the form of the data in Section 2. In Section 3, we show how geodesics fitted to the data can form satisfactory approximations. In Sections 4 and 5, we describe the modelling of the reduced data and give our conclusions in Section 6.

2 Facial Motion Data

Cleft lip and/or palate is one of the most common birth defects. Treatment can involve a sequence of surgeries starting in infancy and extending into the teenage years. The initial surgeries aim to correct the most severe problems while later surgeries further refine the appearance of the face. Naturally, much attention is paid to the aesthetics of the face as this is crucial for social interaction. However, the functionality of the face with respect to speaking, eating and expression is also important, but has received less research attention. There is some concern that surgery designed to improve the static view of the face might damage muscles and nerves and thus damage functionality. See Bardach and Morris (1990), Bardach et al. (1987) and Marsh (1990). Indeed this is a concern with other types of facial surgery.

Trotman et al. (2007) describes an observational clinical trial designed to measure the effects of surgery on the facial soft tissue motion of children with cleft lip/palate. Some of the statistical methods developed for this study and referred to in Trotman et al. (2010) are described in detail in this article. During the trial, three groups of subjects were followed longitudinally for 15 months. One group was comprised of subjects with repaired cleft lip/palate who received lip revision surgery (revision), a second group included subjects with repaired cleft lip/palate who did not have a lip revision, and the third group were non-cleft control subjects. Motion data were collected from all subjects at four time points. For the revision group this data collection occurred at 3 months and just before lip revision surgery, and then again at 3 and 12 months after surgery. Data collection for the non-revision and non-cleft groups occurred at similar times. For each of the four visits, the subjects were instructed to perform six standardised motions - cheek puff, grimace, lip purse, mouth open, natural and maximal smile. Each motion was repeated five times.

The data were collected using motion capture technology. Thirty eight small reflective markers were attached to the face at the locations shown in Figure 1. The 3D locations of the markers were recorded at 1/60 second intervals for four seconds (five seconds for natural smiles). Thus each motion is described by $38 \times 3 \times 240 = 27,630$ numbers. Testing with known rigid objects reveals that there are usually relatively small measurement errors. However, some missing data does arise due to the difficulty in distinguishing markers that are close or other technical problems. Some larger errors are also observed in these circumstances. Another difficulty is that although anatomy helps determine the placement of some of the markers, there is some variation in the placement between subjects and within subjects over successive visits. Indeed, we have some essentially "pseudo-landmarks". An analysis of changes in static shape over time would be sensitive to this variation. However, the adjustment described in Section 4.1 removes much of this variation.

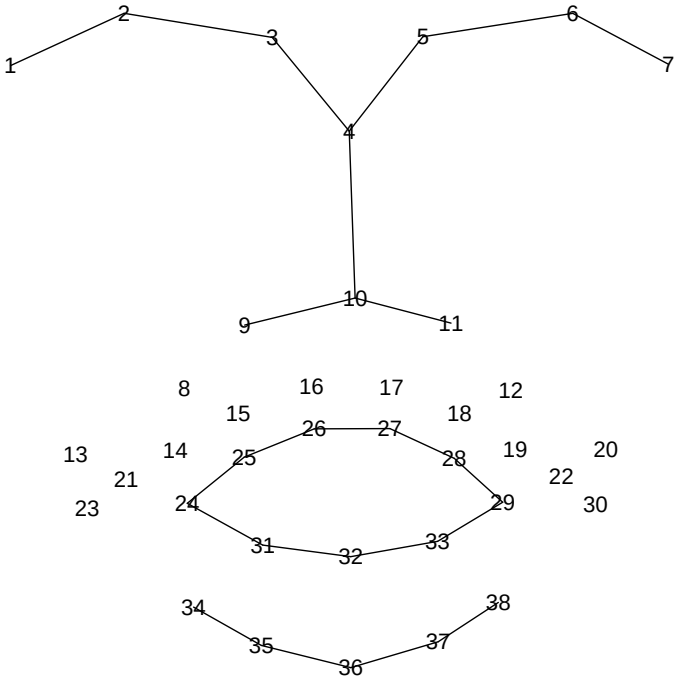

Figure 1: Location of the 38 face markers. The average static face is shown. Lines suggest the location of the eyebrows, nose, lips and chin.

Example movies of facial motions can be found in the supplementary material referred to in the Appendix.

3 Motion on geodesics

A continuous shape change maps out a trajectory in shape space. For example, when $k = 3$, the shape space of triangles can be identified with the surface of a sphere. The simplest form that this trajectory could take is a geodesic, that is the shortest path connecting two points in the shape space. In practice, we would not expect any observed trajectory to follow a geodesic exactly, but

we may wish to check whether a geodesic forms a satisfactory approximation. In Euclidean space, a principal components analysis (PCA) with a dominant first component indicates data that lie close to a straight line, which would be a geodesic in this space. We project shape data onto a tangent space with an origin at the mean shape as described in Dryden and Mardia (1998), p. 76 ff. and justified in detail in Kent and Mardia (2001). The PCA is then performed in the tangent space. Only geodesics in the shape space that pass through the origin will map to straight lines in the tangent space. If the data lie close to a geodesic passing close to the origin of the tangent space projection, the first component will be strongly dominant.

An alternative approach to checking for geodesics is found in Le and Kume (2000) where a multidimensional scaling of the pairwise distance matrix between shapes is performed. Again, a dominant first component indicates that the shapes lie close to a geodesic. We have chosen to use the PCA here because it is faster to compute with the large number of shapes in our examples. In Kenobi et al. (2010), two more sophisticated methods of geodesic fitting based on minimizing the sums of squares of Procrustes distances are presented. The authors remark that their methods will agree closely with PCA for shapes that do not vary much, as is the case with the data of interest here. Kume et al. (2007) presents a spline based method for fitting shape curves beyond geodesics, but this applies only to planar shape data. In Huckemann and Hotz (2009), the geodesics are fitted in the shape space, not the tangent space. Again this applies only to planar data, but there is evidence that fitting the PCA in the tangent space provides good approximations. Kaziska and Srivastava (2007) describes the fitting of curves to 2D gait data. Because the data take the form of human silhouettes, the shape space is different, although related issues arise in the analysis. Another 2D approach to gait modelling that uses statistical ideas about shape can be found in Veeraraghavan et al. (2005) which also contains extensive references to human motion modelling in the computer science literature. Fletcher et al. (2004) develop a method called principal geodesic analysis with application to shapes that are described by boundaries rather than landmarks.

3.1 Closeness to the geodesic for short motions

We computed the PCA on the facial shape motions described earlier using the R **shapes** package of Dryden (2009). Each facial motion $Y(t)$ is composed of f shapes $Y_1, \ldots, Y_f$ where $f = 240$ (except for the natural smiles where $f = 300$). Some preprocessing of the data is necessary. Short gaps of missing values are interpolated while larger gaps result in the motion being discarded. Erroneous measurements due to technological deficiencies are screened out. Some care is important here because PCA for shapes is very sensitive to outliers. Facial motion is smooth and continuous so there should not be any true outliers. Because each PCA is performed on shapes from a single individual, there is no reason to expect a change in size during the motion. Therefore, only the rotation and translation effects were scaled out. We have $Y_i \rightarrow Z_i$ in the tangent space and then PCA on the $f \times 3k$ matrix Z where k is the number of landmarks.

The median percentage variance explained in the first two principal components(PC) are given in Table 1. We see that the first PC is strongly dominant. The dominance of the first PC varies across the type of motion. The first PC was somewhat more dominant for the control subjects than the cleft subjects. Given the large sample size, the difference is statistically significant and may reflect greater irregularity in the motion of cleft subjects. However, no significant longitudinal

change was observed in those cleft subjects who had surgery.

	Cheek Puff	Grimace	Lip Purse	Mouth Open	Natural Smile	Maximal Smile	Control	Cleft
1st PC	82.1	93.6	91.9	98.7	93.6	94.9	94.8	93.0
2nd PC	10.2	3.4	4.3	0.7	3.4	2.9	2.7	3.7

Table 1: Median percentage of variance explained from a PCA of facial motion shape curves.

One might wonder if only the motions in this particular dataset are close to a geodesic so we investigated two other large databases of short duration human motions. In the first, a seated subject sits in a chair and uses the right hand to grasp a block in various locations where the position of the thumb and index finger in grasping the block is also varied. The orientation of the upper body is described by 17 markers and the average sequence of motion has 60 frames lasting 3 seconds. 20 subjects were used in the experiment and 8496 motions were analyzed. The data and motivation behind the experiment are described in more detail in Faraway and Choe (2009). The median percentage variation explained by the first PC was 97.3% while the second PC was 2.1%.

Another experiment involved standing subjects who were asked to place objects with one or two hands onto shelves. The shelves ranged from ankle to shoulder height and were placed in front and to the side of the subject. 20 subjects were used making a total of 7075 motions. The motions were described using 22 markers taking an average of 76 frames which is about 4 seconds of motion. The data and experiment are described in more detail in Faraway (2004). The median percentage variation explained by the first PC was 93.8% while the second PC was 4.8%.

Of course, we do not suggest all shape motion curves are well approximated by geodesics. The motions we have analysed are simple and short in duration. More complex motions over a longer period of time could not be expected to follow a geodesic. However, more complex motion of shapes could be approximated by piecewise geodesics.

3.2 Variation around the geodesic

Although the first component is dominant in these analyses, it is still possible that some interesting structure may lie in the second component. In particular, there may be some commonality in the second PC across different motions. It is difficult to compare the direction of the second PC because each of the PCAs on the individual motions are computed independently. The different coordinate systems generated by these PCAs are then difficult to compare directly. However, we can more easily compare the PC scores. In Figure 2, we see the PC scores for a single subject performing a maximal smile. The first PC score represents the progress along the direction of the first PC. This corresponds to the major direction of motion. The score function demonstrates how the face starts in a rest pose, moves into a smile, holds and then returns to a rest pose. The score functions are quite smooth. One minor problem is that the sign of the score function is arbitrary from the eigendecomposition for the PCA. We have inverted the score function where necessary so that the initial value is always negative.

The second PC score shows no such clear pattern. The functions are still somewhat smooth but are rougher than for the first PC. The curves all start with a negative and end with a positive value

because of the arbitrary inversion standardisation mentioned above and the roughly linear trend. They show no interesting consistent pattern even though the curves come from the same control subject performing the same motion. The plot is representative of the second PC scores that are seen for other subjects. A plot of the first against the second PC score also reveals no particular structure. We conclude that while the motions may diverge from a geodesic, they do so in no consistent way and so we conclude there is just smooth random variation around the geodesic. This random variation is still important as omitting it from simulated or predicted motions will result in unnaturally smooth animations.

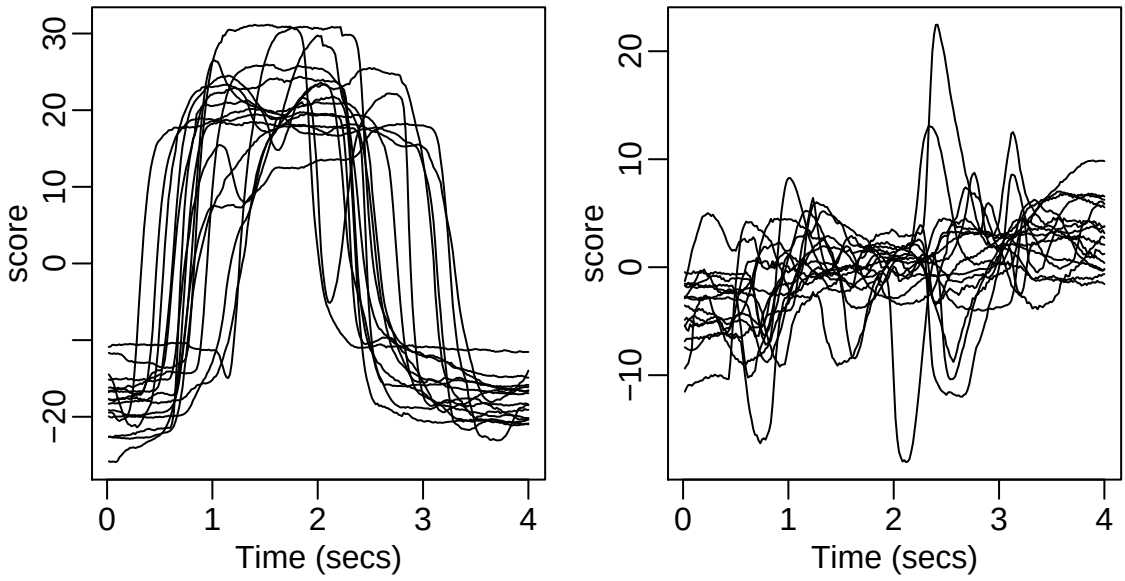

Figure 2: Principal components scores for a control subject. The first plot shows the scores for the first PC plotted against time (frames) while the second plot shows the scores for the second PC.

3.3 Filtering effect of projection to the geodesic

The shape trajectory and geodesic lie in a high dimensional space (over 100 in this example) making it difficult to visualise how they relate in full, but we can observe some lower dimensional projections. In the first panel of Figure 3, we see a plot of the relative change in distance between the lip corners (markers 24 and 29 in Figure 1) for all maximal smiles for the same control subject as in Figure 2 during one of the four visits. The rest position is computed as the median over the first ten frames of motion. We see a common pattern, but there is some local noise as well as larger scale variations from this pattern. Now we project the shape trajectory onto the geodesic and recompute this same relative distance as seen in the second panel of Figure 3. We see that some smoothing has occurred at both the smaller and larger scales. Furthermore, we can see that this smoothing is not what we would expect from directly smoothing the curves in the first panel.

We have filtered out the components of variation due to all but the first PC. By projecting onto the geodesic, we are able to smooth over the whole facial trajectory which is more effective than trying to smooth parts of the motion individually. We also see that projection onto the geodesic still retains substantial variability.

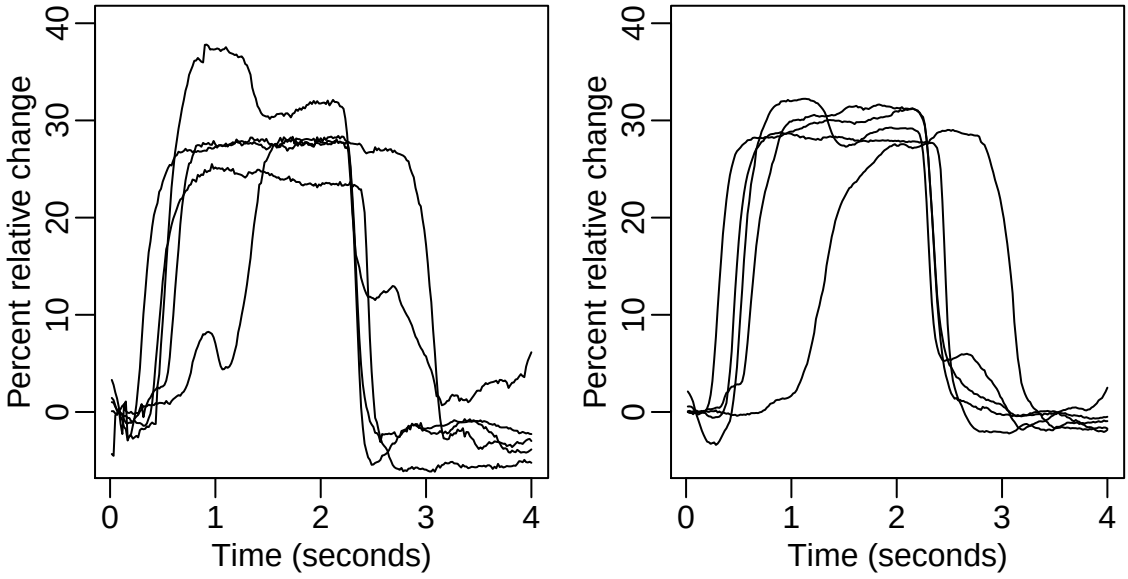

Figure 3: Relative percentage change in the distances between the lip corners over time. Plot on the left shows the observed data. Plot on the right shows the projection onto the geodesic.

3.4 Decomposition of the motion

We can now decompose the motion into two parts. We have the geodesic and the rate of motion along that geodesic which we propose to model independently. The geodesic can be described by any two distinct points along the curve. We will choose the starting pose and the pose when the particular facial expression is maximally realised. We considered two ways of doing this:

1. Let the first PC scores function be $s_1(t)$ and let $t_{max} = \arg\max_t s_1(t)$. We take $Y(0)$ and $Y(t_{max})$ as the starting and extremal poses.

2. Let the first eigenvector of the PCA be e_1, then the fitted geodesic is represented by $\bar{Z} + s_1(t)e_1$ in the tangent space. The projections back to the shape space for $t = 0$ and $t = t_{max}$ represent the starting and extremal poses.

The first uses observed shapes to describe the geodesic while the second uses shapes from the fitted geodesic. We find that the latter method produces less variable results and is used in the analysis to follow.

4 Shape Comparison

Now we consider only the amount of shape change as represented by the starting and extremal pose. The rate of this change is considered in the next section. The reduction of the shape curve $Y(t)$ to just two shapes has greatly simplified the analysis, but some difficulties remain.

4.1 Adjustment for initial pose

When comparing extremal poses for different individuals, we are interested in differences in motion much more than differences in static shape. Individuals naturally have differently shaped faces, but these differences are not immediately interesting for our purposes. We wish to scale out these differences so we adjust all motions to start from a common mean initial shape. For the collection of starting and initial shapes that will be compared, we compute the tangent space at the mean shape. In contrast to the tangent spaces for individual motions, we now scale out the size as well as the translation and rotation because different individuals are being compared. We compute the mean over the starting poses as $\bar{Z}(0)$. In this tangent space:

$$Z_i(t_{max}) \to Z_i(t_{max}) + (\bar{Z}(0) - Z_i(0))$$

which can be projected back to the shape space as needed. This reduces the information for a single motion to one shape, which is the adjusted pose at the extreme. This adjustment will be effective provided there is not too much distortion from the tangent space projection, but this assumption has already been made in the prior analysis.

4.2 Paired comparison of shapes

Consider the set of $Z_i(t_{max})$, $i = 1, \ldots, n$, of initial position-adjusted shapes at the extreme pose. We again make a tangent space projection at the mean shape. This is now the third variety of projection made here. We can average over subgroups in the tangent space and then project back to the shape space.

We will focus on a subset of the full data: the maximal smile motions of the revision group subjects for visits one and four as presented in the Appendix. Readers interested in the substantive conclusions for cleft lip/palate surgery should refer to Trotman et al. (2010). We apply the procedures described above, plotting the mean outcomes in Figure 4. As can be seen, there is almost no difference in mean extremal poses. It is also possible to animate these poses, but this requires an estimation of the relative progress along the geodesics to be considered in the next section.

More formal testing is problematic. A method for comparing two independent groups of shapes is described in Goodall (1991), but we require a paired comparison. In Mardia and Walder (1994), such a method is presented, but is suitable only for 2D data. The repeated measures structure of the data presents a further difficulty as there is little research on modelling shape data of this type. Barry and Bowman (2008) present an approach for linear mixed modelling of 3D shapes that might provide an answer. Unfortunately, the method is only practical for relatively small numbers of landmarks (up to 10 or so) while we have 38 markers. A related problem concerning the symmetry of faces can be found in Bock and Bowman (2006). We will follow a different approach here.

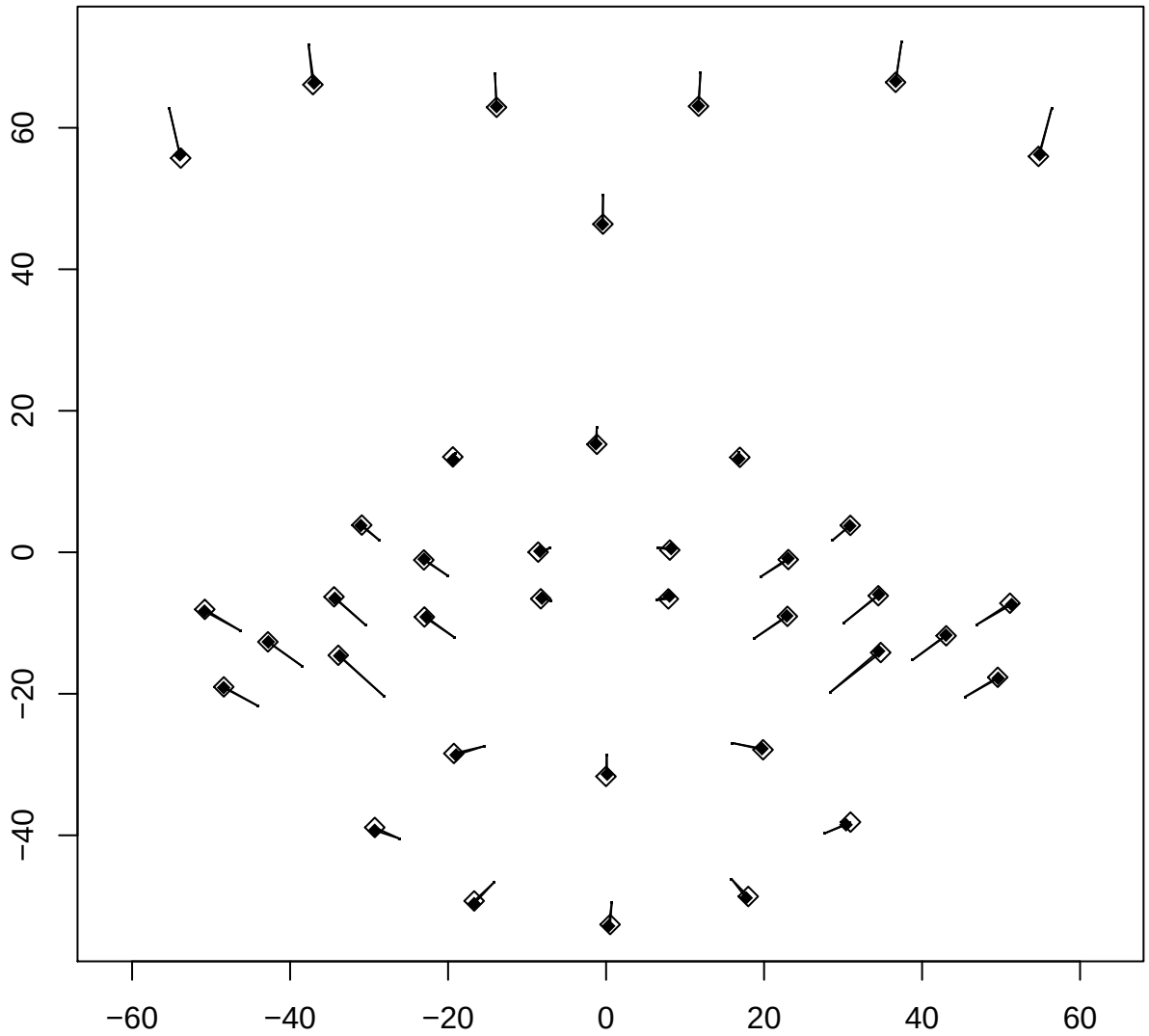

Figure 4: The mean effect of surgery on the maximal smile of patients. The face is viewed from the front. The mean initial pose is marked by the points of the "pins". The mean extremal position before surgery is marked with an open square and the mean position after surgery with a closed square. Lines connecting the extremal positions with the initial positions are drawn.

Although the shape space is high dimensional, we might expect that observed variation lies mostly in a lower dimensional space. We consider the percentage of variation explained by each component of the PCA in the relevant tangent space. For the particular example here, the first five values are 36.5%, 27.0%, 6.1%, 3.6% and 3.3% respectively. The percentage of variation explained by the first component is much lower than that seen in the PCA of the shape curve trajectories. The current PCA is of adjusted extremal poses so there substantial variation between subjects. We consider the first two components as sufficient to represent the variation in the shapes.

After selecting the number of components, each shape is now represented by a short vector of scores $\mathbf{s}$. In the example, we can examine how these scores change after surgery. Let w be a vector of covariate information and c identify the principal component. In our example, w is a binary factor representing before and after surgery while c is a binary factor distinguishing the two principal components. We construct a mixed effects model taking the form:

$$s_i = \beta_0 + \beta_c c_{(i)} + \beta_w w_{(i)} + \beta_{wc} w_{(i)} c_{(i)} + \text{subject}_{(i)} + \text{visit}_{(i)} + \text{component}_{(i)} + \varepsilon_i$$

Suppose that the data are arranged such that a single score value appears in each row with the corresponding coded values of c and w. The parenthesised subscripted i indicates the level of the variable in that row of data. The interaction of c with w allows a different effect for each component of the score. Appropriate generalization is necessary for c with more levels or multivariate w. The term $\text{subject}_{(i)} \sim N(0, \sigma_s^2)$ is a random effect representing the subject while $\text{visit}_{(i)} \sim N(0, \sigma_v^2)$ represents the nested visit effect. For the component nested within the visit effect, we have $\text{component}_{(i)} \sim N(0, \sigma_c^2)$. The final term $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ represents the lowest level of variation from motion to motion. This is a relatively simple model containing all elements, but some elaboration might be considered to allow unequal variances or correlations structures for the random effects. This model was fitted using the **nlme** R package of Pinheiro and Bates (2000). A comparison to a model without w to test for the effect of surgery yielded a $p-$value of 0.92, confirming earlier impressions of little difference due to surgery.

The random components estimates are $\hat{\sigma}_s = 3.9$, $\hat{\sigma}_v = 0.0$, $\hat{\sigma}_c = 8.7$ and $\hat{\sigma}_\varepsilon = 5.2$ where only the relative size of these values is of much interest. There are only two levels of visit and component in this example, and so these two constituents of the variability are difficult to distinguish. The larger size of the latter relative to the former should not be taken at face value. This estimated effect for surgery is 0.3 which is much smaller than the estimated natural variation. The relative size of the random effect components demonstrates the variability in motion from one visit to the next and the necessity for substantial replication in such studies.

4.3 Individual changes

In Figure 5, we show the change in the scores averaged within subjects. As can be seen, the average change is around zero, but some subjects do show considerable variation. We estimate the standard error of these means as 9.0. There is some evidence of a change in at least two of the subjects which are labelled in the plot.

The differences for the two subjects are shown in Figure 6. For the first subject, the difference is that before surgery, the subject made a closed mouth smile while after surgery the subject smiled

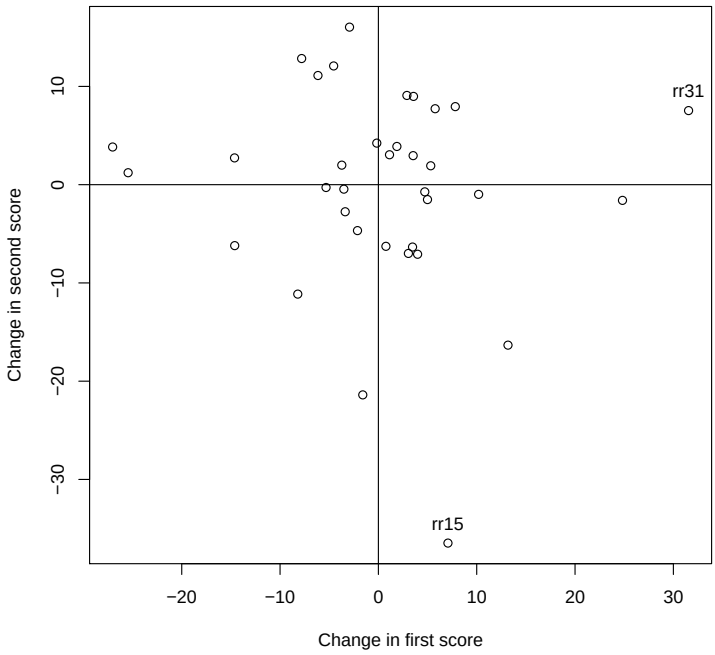

Figure 5: Change in mean principal component scores by subject due to surgery. Subjects showing some evidence of change are labelled.

with an open mouth. This kind of variation is also observed in control subjects and so it likely has nothing to do with surgery. In the second subject, the difference is more in the magnitude of the smile, with a larger smile after surgery. Again, this kind of variation is seen in control subjects, but it is plausible that surgery might affect this. The analysis is not intended to make definite decisions, but to point to cases where further expert judgement may be required

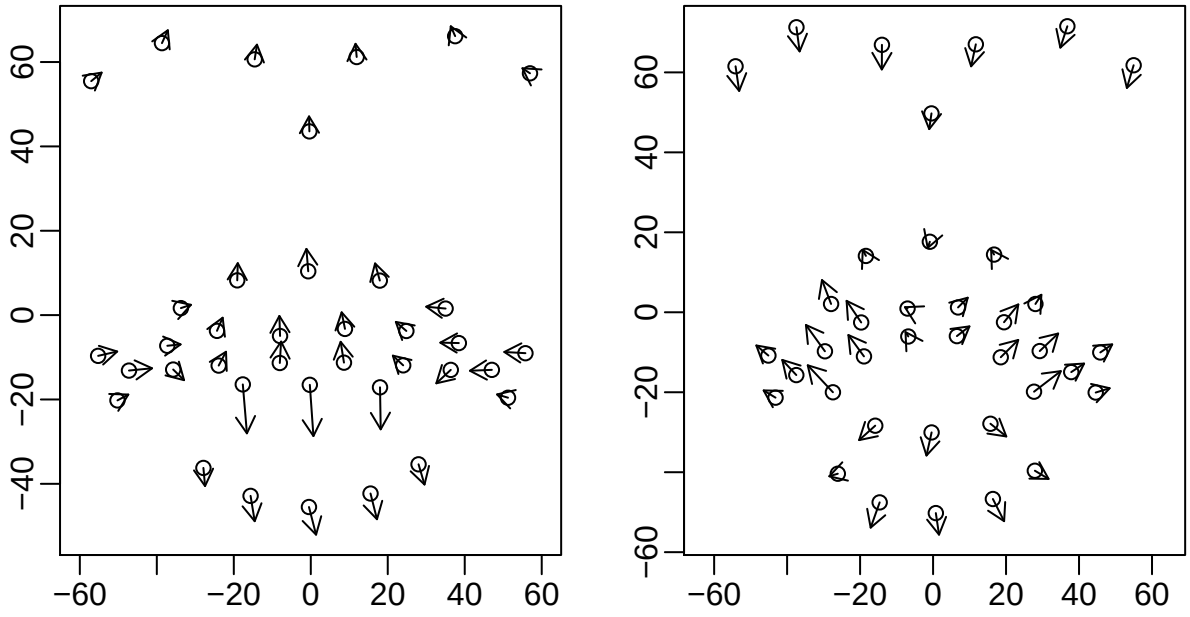

Figure 6: Change in pose before and after surgery for subject RR31 on the left and RR15 on the right. Mean before pose is denoted by an open circle with the arrow ending at the after surgery mean pose.

5 Score Profiles

An examination of the first PC scores in Figure 2 shows that the rate of motion can be readily described by the four transition times marking the move from rest to the expression and back again. Although more sophisticated methods for curve alignment are available such as found in Ramsay and Li (1998), we found that a simpler ad hoc method produced subjectively more effective and robust results.

We set the starting value as the median of the first ten observations and the ending value as the median of the last ten. We compute the maximum and the range. We set the first cutpoint as the last time before the maximum that the curve is within 5% of the range above the starting value. The second point is the first time the curve is within 10% of the range below the maximum. The third and fourth cutpoints are defined symmetrically.

We define three times of interest. The first is the reaction time, that is the time to the first cutpoint. The second is the move time, that is the time between the first and second cutpoints. The third is the hold time, that is the time between the second and third cutpoints. We compared these times to the scores from the PCA of the adjusted extremal positions as computed in the previous section. We saw no relationship, reinforcing the assumption that the profiles and the extremal poses can be treated independently.

We can test for an effect due to surgery on these times with mixed effect models of the form:

$$t_i = \beta_0 + \beta_w w_{(i)} + \text{subject}_{(i)} + \text{visit}_{(i)} + \varepsilon_i$$

with random effects defined as before. This model reveals no significant effect for any of the three times of interest.

Furthermore, a comparison of these times t_i with scores s_i in the previous section shows no evidence of a dependence between progress along the geodesic and the direction of that geodesic. This justifies the independent treatment of these two constituents of the motion.

Given that there seems to be little distinction between the score profiles, we may reasonably average them to produce a profile for constructing predicted motions. We map a score profile $s(t) \to s(h(t))$ where $h(t)$ is a piecewise linear function that maps the cutpoints for this score to the mean cutpoints across the data. Now given an initial pose and an extremal pose, we can construct the geodesic using this mean score profile for the rate of progress. The mean motion over the data constructed in this manner is shown as a movie as described in the appendix.

6 Conclusions

The conclusions regarding the cleft lip data are described in detail in Trotman et al. (2010). Surgery is often motivated by a desire to improve the static appearance of the face. At the same time, the appearance of the face during movement or function is equally, if not more, important. We found no significant evidence that surgery had an effect on average on the facial function of patients. However, there was evidence that the function of some individual patients improved and some deteriorated.

Continuous shape change data is naturally high-dimensional. Interpretable data analysis requires the identification of the low-dimensional structure in the data. For the data presented here, the shape change occured mainly along a geodesic in the shape space. We have also shown how this approximate geodesic feature can be seen in other datasets involving human motion. When considering shape changes in collections of objects, it is also important to distinguish between the static or initial shape and the manner in which it changes. We have demonstrated how the shape change can be considered independently. Further dimension reduction in the description of these shape changes was then possible to the extent that a single instance of shape change (a motion in our example) could be reasonably described with just two numbers. Once this reduction is made, we may apply well-known statistical methods for one or low-dimensional responses.

The methods described here are more likely to work for data where the changes in shape are not extensive. For more complex shape change data, such as the motion of a dancer for example, geodesic approximation may not suffice. In Faraway et al. (2007) a method of modelling 3D

trajectories relying on a small number of Bézier control points is presented. This method could be extended to modeling the higher dimensional curves in the tangent space for such shape curves. With a small number of control points, the subsequent parametric mixed effect models used in this article would still be tractable.

Acknowledgments

The research was supported in part by grant DE13814 from the US National Institute for Dental Research.

Appendix

Supporting materials may be found at `www.bath.ac.uk/∼jjf23/face/shapcurve/`. The contents are:

1. Motions from the subset of data analyzed in detail in the article.
2. Animated motions including:
   (a) Raw data
   (b) Mean motions

References

Bardach, J. and H. Morris (1990). *Multidisciplinary management of cleft lip and palate*. Philadelphia: WB Saunders Company.

Bardach, J., K. Salyer, and I. Jackson (1987). *Surgical techniques in cleft lip and palate*. Year Book Medical Publishers.

Barry, S. and A. Bowman (2008). Linear mixed models for longitudinal shape data with applications to facial modeling. *Biostatistics 9*(3), 1–11.

Bock, M. and A. Bowman (2006). On the measurement and analysis of asymmetry with applications to facial modelling. *Journal of the Royal Statistical Society-Series C Applied Statistics 55*(1), 77–92.

Bookstein, F. (1991). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press.

Dryden, I. (2009). *shapes: Statistical shape analysis*. R package version 1.1-3.

Dryden, I. and K. Mardia (1998). *Statistical Shape Analysis*. Chichester: Wiley.

Faraway, J. (2004). Human animation using nonparametric regression. *Journal of Graphical and Computational Statistics 13*, 537–553.

Faraway, J. and S. B. Choe (2009). Modeling orientation trajectories. *Statistical Modelling 9*, 51–68.

Faraway, J., M. Reed, and J. Wang (2007). Modeling 3D trajectories using Bézier curves with application to hand motion. *Applied Statistics 56*, 571–585.

Fletcher, P., C. Lu, S. Pizer, and S. Joshi (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions On Medical Imaging 23*, 995–1005.

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological) 53*, 285–339.

Huckemann, S. and T. Hotz (2009). Principal component geodesics for planar shape spaces. *Journal of Multivariate Analysis 100*(4), 699–714.

Kaziska, D. and A. Srivastava (2007). Gait-Based Human Recognition by Classification of Cyclostationary Processes on Nonlinear Shape Manifolds. *Journal of the American Statistical Association 102*(480), 1114–1124.

Kendall, D., D. Barden, T. Carne, and H. Le (1999). *Shape and Shape Theory*. Chichester: John Wiley.

Kenobi, K., I. Dryden, and H. Le (2010). Shape curves and geodesic modelling. *Biometrika 97*, 567–584.

Kent, J. and K. Mardia (2001). Shape, Procrustes tangent projections and bilateral symmetry. *Biometrika 88*(2), 469–485.

Kume, A., I. Dryden, and H. Le (2007). Shape-space smoothing splines for planar landmark data. *Biometrika 94*(3), 513–528.

Le, H. and A. Kume (2000). Detection of shape changes in biological features. *Journal of Microscopy 200*(2), 140–147.

Mardia, K. and A. Walder (1994). Size-and-shape distributions for paired landmark data. *Advances in Applied Probability 26*(4), 893–905.

Marsh, J. (1990). When is enough enough? Secondary surgery for cleft lip and palate patients. *Clinics in plastic surgery 17*(1), 37.

Pinheiro, J. C. and D. M. Bates (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.

Ramsay, J. and X. Li (1998). Curve registration. *Journal of the Royal Statistical Society, Series B 60*, 351–363.

Small, C. (1996). *The Statistical Theory of Shape*. New York: Springer.

Trotman, C.-A., J. Faraway, C. Philips, and J. van Aalst (2010). Effects of lip revision surgery in cleft lip/palate patients. *Journal of Dental Research 89*, 728–732.

Trotman, C.-A., C. Phillips, G. Essick, J. Faraway, S. Barlow, J. van Aalst, and L. Rogers (2007). Functional outcomes of cleft lip surgery. Part I: Study design and surgeon ratings of lip disability and need for lip revision. *The Cleft Palate-Craniofacial Journal 44*(6), 598–606.

Veeraraghavan, A., A. Roy-Chowdhury, and R. Chellappa (2005). Matching shape sequences in video with applications in human movement analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27*(12), 1896–1909.